FISH & RICHARDSON P.C.

601 Thirteenth Street N.W.
Washington, DC 20005

Telephone
202 783-5070

Facsimile
202 783-2331

Web Site
www.fr.com

Frederick P. Fish
1855-1930

W.K. Richardson
1859-1951

BOSTON
DELAWARE
NEW YORK
SILICON VALLEY
SOUTHERN CALIFORNIA
TWIN CITIES
WASHINGTON, DC

JC639 U.S. PTO
12/28/99

jc511 U.S. PTO
09/472964
12/28/99

December 28, 1999

Attorney Docket No.: 06105-003002

Box Patent Application
Assistant Commissioner for Patents
Washington, DC 20231

Presented for filing is a new continuation patent application of:

Applicant: DAVID K. GIFFORD

Title: REPLICA ROUTING

Enclosed are the following papers, including those required to receive a filing date under 37 CFR 1.53(b):

	Pages
Specification	20
Claims	7
Abstract	1
Declaration	1
Drawing(s)	9

Enclosures:

— Copies of Small Entity Declaration, which were filed in the parent application. Please enter these in this application. This application is entitled to small entity status.
— Postcard.

This application is a continuation (and claims the benefit of priority under 35 USC 120) of U.S. application serial no. 08/779,770, filed January 7, 1997. The disclosure of the prior application is considered part of (and is incorporated by reference in) the disclosure of this application.

Basic filing fee	$380
Total claims in excess of 20 times $9	$0
Independent claims in excess of 3 times $39	$39
Fee for multiple dependent claims	$0
Total filing fee:	$419

A check for the filing fee is enclosed. Please apply any other required fees or any credits to deposit account 06-1050, referencing the attorney docket number shown above.

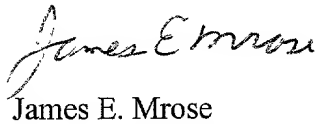
If this application is found to be incomplete, or if a telephone conference would otherwise be helpful, please call the undersigned at (202) 783-5070.

Kindly acknowledge receipt of this application by returning the enclosed postcard.

Please send all correspondence to:

JAMES E. MROSE
Fish & Richardson P.C.
601 Thirteenth Street, NW
Washington, DC 20005

Respectfully submitted,



James E. Mrose
Reg. No. 33,264
Enclosures
JEM/adt
40019411.doc

Applicant or Patentee: David K. Gifford
Serial or Patent No.: 08/779,770
Filed or Issued: January 7, 1997
For: REPLICA ROUTING

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR 1.9(f) and 1.27(b)) - INDEPENDENT INVENTOR

As a below named inventor, I hereby declare that I qualify as an independent inventor as defined in 37 CFR 1.9(c) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code, to the Patent and Trademark Office with regard to the invention entitled REPLICA ROUTING described in

[ ] the specification filed herewith.
[X] application Serial No. 08/779,770, filed January 7, 1997.
[ ] patent no. , issued .

I have not assigned, granted, conveyed or licensed and am under no obligation under contract or law to assign, grant, convey or license, any rights in the invention to any person who could not be classified as an independent inventor under 37 CFR 1.9(c) if that person had made the invention, or to any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern or organization to which I have assigned, granted, conveyed, or licensed or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

[X] no such person, concern, or organization.
[ ] persons, concerns or organizations listed below*.

*NOTE: Separate verified statements are required from each named person, concern, or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Full Name: _____

Address: _____

[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Inventor: David K. Gifford

Signature: _____ Date: 5/12/97

APPLICATION

FOR

UNITED STATES LETTERS PATENT

06105-003002 - 4253806

TITLE: REPLICA ROUTING

APPLICANT: DAVID K. GIFFORD

REPLICA ROUTING

Background of the Invention

This invention relates in general to an internetwork replica routing system and more particularly relates to a system for directing a client computer to a server replica that is estimated to provide good performance for the client computer.

The recent rapid growth of users of international public packet-switched computer internetworks such as the Internet has created a large demand for the information and services they contain. The replication of services in an internetwork makes it possible for such services to service many users.

Certain known approaches for automatically directing client computers to servers include, for example, round robin DNS and loading balancing DNS, which direct users to one of a number of server replicas in an attempt to balance the load among many servers. In another approach called multiple hostnames, content is spread over multiple servers, each with a separate hostname. Web pages returned to users contain links that point at the replica that has been selected for the user based on load-balancing concerns and replica content considerations. In another approach called Internet load balancing, a hardware component automatically distributes user requests sent to a single IP address to one of a number of server replicas to implement load balancing. Another approach is resonate dispatch that combines load balancing with replica capability to automatically direct users to a replica that is operational, is not overloaded with requests, and contains the requested information.

Summary of the Invention

The invention provides a network server replication system that uses a new method called replica routing to

automatically direct a client computer to a server replica that will perform well for the client given the location of the client in the internetwork. More specifically, client computers contact a replica router that transparently redirects them to a server replica in the internetwork that will perform well given the client's network location and the estimated performance of the internetwork.

The replica routing provided by the invention allows a client computer to access key replicated information near the location of its use in the internetwork. In particular, once an appropriate server replica is established, replica routing according to the invention automatically directs clients that are "nearby" in the internetwork to that particular replica, thereby providing high-performance access to the replicated server. For example, the invention will enable high-performance access to network applications, such as video, that are highly sensitive to the delay and bandwidth that are introduced by network components such as network links and network routers. This high-performance access is especially important given that public computer internetworks typically include switching, transmission, and host computer components controlled by many individuals and organizations.

In certain embodiments each server replica creates a replica advertisement that summarizes information about its location in the internetwork and its observation of the local internetwork topology and performance. A server replica automatically passes its replica advertisement into the replica routing system. The replica router relies upon replica advertisements supplied by each server replica and other optional measurements to route a client computer to one or more server replicas. New replicas can be flexibly added without undue administrative overhead.

As the number of server replica grows it may become

impractical or inefficient for every replica router for a service to contain the replica advertisements for all server replicas. One particular method for replica routing according to the invention allows replica routers to be optionally arranged in a hierarchy, and for replica advertisements to be propagated only part way up the replica router hierarchy. During the replica routing process client requests are automatically sent down the hierarchy until they reach a replica router that is sufficiently knowledgeable about a replica's internetwork location to make an informed replica routing judgment. Thus, not all of the replica advertisements for a particular service have to be contained in a single replica routing server.

A second reason for introducing a hierarchy of replica routers is for security concerns. Since a replica advertisement can contain sensitive information about internetwork characteristics and topology, an organization can choose to create a private replica router inside of a private internetwork (an intranet) to contain this information. In one particular embodiment of the invention, client requests from inside of the intranet will be automatically directed to this private replica router, while client requests from outside of the intranet will use other replica routers that do not need to know the detailed advertisements for replicas they cannot access.

In another aspect of the invention, a client applet can assist in the replica routing process. The client applet can determine certain characteristics of the client internetwork environment, and send these to the replica router as additional information to aid the routing process. The replica router can return more than one replica address to the client applet, and the client applet can then perform empirical performance experiments to choose the best server

replica for the use of the client.

Numerous other objects, features, and advantages of the invention will appear from the following description taken together with the drawings.

Brief Description of the Drawings

Fig. 1 is a block diagram of a replica routing system according to the invention.

Fig. 2 is a diagram of a replica routing hierarchy.

Fig. 3a-3c is a flow chart illustrating the creation of a replica advertisement and the processing of a replica advertisement by replica routers.

Fig. 4a-4b is a flow chart illustrating the processing of a client request to a replica router that results in the client being redirected to another replica router or a replica.

Fig. 5a-5b is a flow chart illustrating the alternative processing of a client request to a replica router with the aid of a client replica routing applet.

Fig. 6 is a flow chart illustrating a process for determining internetwork performance in a neighborhood of adjustable size.

Detailed Description

A replica routing system 100 as shown in Fig. 1 employs public internetwork 10 and intranets 11 and 12 protected by firewalls 20 and 21 to interconnect a plurality of client computers 31, 32, 33, and 34, server replicas 41, 42, 43, and 44, replica routers 51 and 52, and master servers 61 and 62. Each replica router 51 and 52 has a corresponding replica routing database 71 and 72. A user of the system employs a client computer 31, 32, 33, 34 to access services provided by master servers 61, 62 and server replicas 41, 42, 43, 44, and is directed to an appropriate server replica by one or more replica routers 51, 52. Client computers can

include user workstations and proxy servers. Master servers 61, 62 are used to service certain update requests that are not processed at server replicas 41, 42, 43, 44, such as purchasing of goods or user registration that requires the synchronized updating of shared databases. Replica routers, server replicas, and master servers can be implemented on separate computers as shown, or can share computers. For example, a replica router, a server replica, and a master server can exist on the same computer.

The contents of server replicas 41, 42, 43, 44 can be dynamically maintained by a network-based replication method, such as a master-slave scheme or weighted voting, or replicas can be updated by digital broadcast either over the network or by separate multicast or broadcast channels such as satellite or terrestrial links. Alternatively, replicas can either be partially or totally implemented by media that can be physically distributed such as optical disk.

The software architecture underlying the particular preferred embodiment is based upon the hypertext conventions of the World Wide Web. The Hypertext Markup Language (HTML) document format is used to represent documents and forms, and the Hypertext Transfer Protocol (HTTP) is used between client, replica router, server replica, and master server computers. Documents are named with Uniform Resource Locators (URLs) in the network of computers. A document can be any type of digital data broadly construed, such as multimedia documents that include text, audio, and video, and documents that contain programs. In particular, Java applets and ActiveX controls can be contained in or referenced by documents that allow the capabilities of client computers to be automatically extended by the downloading of new programs.

In addition to documents, server replicas can be used to replicate any type of data, including relational databases,

multimedia data, video files, and groupware data. To support access to these datatypes server replicas and master servers can support a variety of standardized protocols in addition to or instead of HTTP, such as standard remote procedure call protocols, database interfaces such as ODBC, and Microsoft's Distributed Common Object Model (DCOM) and their successors.

Server replicas can also replicate programs that are used to generate dynamic content. When a server replica receives from a client and processes an HTTP request for the URL of such a program, the program is executed, and the program produces dynamic content that is returned to the client. HTTP POST data can also be sent to dynamically executed programs using the same URL-based mechanism. Thus, server replicas can be used to produce dynamic content and process client data as well as serve static content. This program replication and dynamic program invocation mechanism can also be employed with other standardized communication protocols.

Server replicas can optionally process certain update and transaction requests and relay them to other server replicas or master servers. For example, a shared project page could be updated at one server replica, and this update could be automatically propagated to all other replicas. Alternatively, form submissions for hard-good orders could be spooled to a file, by a server replica, and then sent at predetermined intervals to a master server for further processing.

The network architecture underlying one particular embodiment is TCP/IP and the family of IP routing procedures. Background on this network technology can be found in TCP/IP Illustrated by Stevens (1994, Addison-Wesley, Reading, Massachusetts), Routing in the Internet by Huitema (1995, Prentice Hall, Englewood Cliffs, New Jersey), and "End-to-End

Routing Behavior in the Internet" by Paxson (1996, SIGCOMM '96 8/96 CA, USA). The term network is used herein to refer to both networks and subnetworks. The term network number or network identifier is used to refer to the IP address of a network including both its network and subnetwork components.

Fig. 2 shows an example hierarchy 200 of replica routers, with router 201 being a root replica router, and with router 203 being a leaf replica router that contains replica advertisements for server replicas in its network neighborhood. More than one replica router can exist at each level of the hierarchy, and there can be multiple root replica routers. The IP addresses of the root replica routers are bound to the DNS name of the service, such as "www.pathfinder.com."

Before discussing how replica routers operate to direct client computers to server replicas that provide good performance for the client computers, this discussion will describe how a client computer or server replica can "discover" its local internetwork topology. Then this discussion will describe how this technique is used in connection with the replica routing system of the present invention.

Fig. 6 is a flowchart for the discovery of local internetwork topology and performance. All routing protocols that are in use on the internetwork are employed, including the Routing Information Protocol (RIP), External Gateway Protocol (EGP), Boarder Gateway Protocol (BGP), Open Shortest Path First (OPSF), Classless Interdomain Routing (CDIR), and their descendants and follow-ons. At step 905 a client computer (or a server replica) sends a network router solicitation message on all connected networks by broadcast or multicast to learn of nearby network routers (standard network routers are not to be confused with the replica routers

according to the invention).

At step 910, network routing table request messages are sent to all of the network routers discovered in step 905, along with any well-known or preconfigured network routers. Responses (routing tables) from the network routers are received by the client computer at step 915. At step 920 the client computer derives from the routing tables the expected performance from the client's network to all of the networks specified in the received routing tables and records this information in a network performance table. The network performance table is a list of rows, in which each row contains a network number, a net mask, and an estimate of the performance from the client to the network number (e.g., an estimate of bandwidth). A net mask (sometimes called a subnet mask) specifies which portions of an IP address contain network and subnetwork identifiers and thus should be matched to a second IP address to determine whether the two addresses are on the same network. Each network performance table entry also includes the net mask for the destination network as reported by the routing table. If no net mask is reported by a network router in a destination network, then a default net mask based on the class of the destination network's IP address (which is inferred from the initial digits of the address) is used, or another pre-specified set of rules is used. If more than one network router offers a route to a distant network, the client computer records only the best-performing route in the performance table. A single metric for replica routing "performance" is used, such as estimated bandwidth. For example, if a particular RIP network router reports the number of network hops it requires to reach a distant network, rather than the estimated bandwidth required to reach the distant network, the number of hops can be converted to estimated bandwidth by simply reducing bandwidth

from an ideal fixed maximum by a fixed amount for each hop reported. Alternatively, if the address of a router on a distant network is discovered in the information received at step 915, it can be "pinged" to attempt to estimate the network performance from the client to the distant network. At step 925 if a configuration-set maximum number of iterations has not been exceeded, then at step 935 all of the network routers that were named in the routing tables received at step 915 that were not previously explored are assembled, and this set of new routers is used at step 910 to learn more about the network neighborhood. Otherwise, at step 930, internetwork performance discovery is completed, yielding a network performance table that is a list of rows, in which each row contains a network number, a net mask, and an estimate of the performance to that network number.

In an alternative embodiment, the maximum depth (maximum number of iterations) explored for a given network router can depend on the network router (e.g., well-known network routers can have a greater maximum depth). In another alternative embodiment, more than one network performance metric can be utilized (such as bandwidth and latency).

Multiple types of network numbers can be used simultaneously in a network performance table, and thus multiple types of network numbers can be used in replica routing databases, including IP network numbers, IPng (next generation) numbers, and their successors.

In an alternative implementation, the network performance map is extended by using a traceroute utility to perform traceroutes to preconfigured IP addresses and to the IP addresses of server replicas. Traceroute utilities are described in TCP/IP Illustrated, Vol. 1, Chapter 8: "Traceroute Program," Stevens (1994, Addison-Wesley, Reading, Massachusetts). Server replica addresses can be discovered by

contacting root replica routers and other replica routers and asking them with a specialized request to transmit their list of server replica and replica router addresses. The use of traceroute can uncover information that is not available from routing table inquiries.

Fig. 3a-3c is a flowchart that describes the operation of server replicas and replica routers, including the creation of replica advertisements. A key concept in the operation of this system is that server replicas and replica routers both perform the functions of selecting parent replica routers and sending "advertisements" to the parent replica routers that include information concerning the server replica's or replica router's address in the internetwork as well as local internetwork topology and performance estimates derived from server replicas or replica routers.

At step 301, if a particular computer is a server replica, control is transferred to step 302, and if it is a replica router, control is transferred to step 306.

At step 302 a server replica creates a network performance table using the method presented in Fig. 6. At step 305 the server replica creates a replica summary record that has one entry for each network it can reach. Each replica summary record entry contains: a network number, the network's net mask from the network performance table (see Fig. 6), an estimate of the performance to the network from the network performance table, and the current time as a timestamp. The entire replica summary record is marked as being created by a server replica.

Alternatively, if a particular computer is a replica router, then at step 306 the replica router scans its replica routing database and deletes any replica summary record entries that have a timestamp that is older than a configuration-set time limit. At step 307 a test is made to

see if this replica router is configured as a root replica router. If so, then control is transferred to step 370 and the root replica router does not create an advertisement (because the root replica router has no parent to which it can send an advertisement). If this replica router is not a root replica router, then at step 310 the replica routing database is used to create a new replica summary record that has multiple entries, one for each network number advertised in a replica summary record in the replica routing database. Each entry in the newly created replica summary record includes: a network number, the net mask for that network number, the best performance metric value for that network number that is advertised in a replica summary record by any server replica or replica router, and the timestamp from this best performing entry in the routing database. The newly created replica summary record is marked as being created by a replica router.

At step 315 logic common to the replica routers and server replicas begins, and the new replica summary record can be modified according to operator-specific rules that are specific to the replica router or server replica. Arbitrary alterations to the new replica summary record can be specified, including: the removal of certain networks; the addition of network numbers with specified network masks, performance metric values, and timestamps that can include a "do not expire value"; manual override, by network number, of network masks and performance metric values or replica summary record entries; and removal of replica summary record entries that do not achieve a specified performance metric value. In this way the operator of a server can ensure that the server serves its intended audience, for example by adding intranet network numbers that cannot be seen from outside the intranet's firewall. Next, the replica router or server replica selects a set of parent replica routers. The

addresses of the parent replica routers are initialized by looking up the replica routers bound to the service's DNS name (such as "www.pathfinder.com"). Alternatively, the set of parent replica routers can be manually configured for more involved hierarchies.

In an alternative embodiment, server replicas contain the same basic information, but are specialized with local features such as having their content in a foreign language or having buyer-location-specific prices. In this embodiment, a server replica's advertisement is altered in 315 as described above to offer to service network numbers that are in a geographical area, e.g., a country, or administrative domain, e.g., a company, regardless of whether these network numbers are actually close to the server replica, to cause client routing to server replicas to be based on content specialization. In particular, the network numbers that the server replica wishes to service can be added to the network numbers ordinarily advertised as being close to the server replica's location (if the network numbers that the server replica wishes to service are different from the numbers advertised as being close to the server replica's location), or the numbers ordinarily advertised are deleted. For example, a server replica having content in the French language could offer to service network numbers that are in France. In this example, when multiple servers offer to service networks in France, then performance metric values will be used to choose the best one for a client request.

Once the set of parent replica routers is determined their addresses are stably recorded for later use.

At step 325 a replica advertisement is constructed that includes the replica summary record, the IP address of the local computer, and the current time. A digital signature is added to the replica advertisement at 330 that is based

upon a service-specific private key known to all replicas and replica routers of a service, and the completed replica advertisement is sent to the parent replica routers in HTTP POST message 335.

After a parent replica router receives the replica advertisement message 335, at step 340 it authenticates the replica advertisement using the public key of the service. Once the advertisement is authenticated, at step 341 a check is made to ensure that the IP address in the advertisement is the same as the source IP address in the header of message 335. If the IP addresses match, control continues at 345, otherwise control continues at 342.

At step 342, if the IP addresses do not match, it means that the replica advertisement has traveled through a firewall (see Fig. 1). The multiple-entry replica summary record in the replica advertisement has a single entry added that includes: the source IP address in the header of message 335, a net mask of all bits "1," a default network metric value, and the current time. This additional entry is added to the summary because the added IP address will be identical to the IP addresses of requests made by clients from behind the same firewall, and thus will match the IP addresses of these client requests. This new replica summary record is marked as having been created by a replica router.

At step 345 the replica advertisement is added to the local routing database at the parent replica router if the advertisement has a more recent timestamp than a previous replica advertisement in the routing database from the same IP address specified in the replica advertisement. Replica advertisements that are superseded by newer advertisements are deleted.

An acknowledgment message is constructed at step 355 that contains the IP address contained in the advertisement,

the timestamp, and a digital signature using the service private key. The acknowledgment message 360 is authenticated by the sending server replica or replica router at step 365, and a timer is set at step 370 to refresh the registration information at a configuration-determined interval. When the timer expires, control of the server replica or replica router returns to step 301.

In an alternative embodiment, replica routers regularly "ping" the servers that are described by replica advertisements in their replica routing databases. Servers that do not respond have their advertisements removed unless their replica summary record contains an entry that has a net mask consisting entirely of "1"s, which indicates that the server is behind a firewall.

Fig. 4a-4b shows a flowchart that describes the process of forwarding a client computer to a server replica using HTTP redirects. At step 505 a user selects an entry from a previously received HTTP page that contains a URL that refers to a replicated service. At step 510 a network request 515 is created, such as a GET or a POST, and request 515 is sent to a root replica router for the service. The IP address of the root replica router is derived from the DNS name or address in the URL selected in step 505.

At step 535 the replica router matches all of the replica summary record entries in the replica routing database to the source IP address in message 515. Address matching of each replica summary record entry is performed based on the portion of each address that constitutes the network identifier according to the net mask contained in the entry. If no matches are found, control is transferred to step 562. If matches are found, at step 545 the N matching replica summary record entries that contain the best network performance metric values are selected and sorted by

decreasing network performance metric value, and the IP addresses contained in the corresponding replica advertisement entries are made the candidate target IP addresses. Each entry in the candidate target IP address list includes a descriptor indicating whether it is a replica router or server replica (this information is determined from the entry's replica advertisement). The number N is a configuration parameter. Control then transfers to step 590.

At step 562 the replica router determines the network route and hop-by-hop delay to the client IP address in the header of message 515 using a utility such as traceroute. If the replica router already has the routing and performance information because of a previous execution of step 562 to the same client address it uses this information if the information is not older than a configuration-set maximum time.

At step 565 the IP address of each network router in the network route to the client is looked up in the replica summary records in the replica routing database, starting at the network router closest to the client. Matching is performed using the net mask in each replica summary record entry. If there are no matching entries in the replica routing database, then at step 575 a default set of pre-specified server replicas is made the set of candidate target IP addresses, and all of the default replicas are marked as server replicas. Control is then transferred to step 590.

If there are matching entries in the replica summary records in the routing database then at step 570 each matching replica summary record entry has its advertised network performance added to the network performance estimated from the client to the network router it matched. One way to estimate this performance is to take the round-trip performance observed from the replica router to the client and

adjust for the round-trip performance from the replica router to the network router that matched. The N matching replica advertisement entries that contain the best aggregate network performance metric values are selected, and the IP addresses contained in these replica advertisement entries are made the candidate target IP addresses. Each entry in the candidate target IP address list includes a descriptor indicating if it is a replica router or server replica; this information is determined from the entry's replica advertisement. The N IP addresses are ordered by decreasing network metric merit. Control then transfers to step 590.

At step 590 a new URL is computed that consists of the URL sent in message 515 with its network address portion replaced by the IP address that is the highest on the candidate target IP address list. The new URL is sent back to the client in redirect message 595.

At step 596 the client processes the redirect. If the new URL points to a replica router, then the client will automatically start again at step 510 using a different replica router. One application of such a redirect is to redirect a client to a replica router that is behind a firewall that is specialized for server replicas in the client's intranet. If the new URL points to a server replica, then the server replica will return pages that contain relative links for all requests that can be serviced from the replica, and absolute links to a master server for all requests that need to be serviced by a master server. Relative links allow the client to carry the local server replica host name from request to request, as well as optional information such as a session identifier or digital receipt. Absolute links created by a server replica can encode similar information including session identifiers, and also always encode the IP address of the referring server replica, so that

the master server can learn the IP address of the referring replica to enable the master server to redirect the client back to the referring replica once the master server has finished its specialized processing.

In an alternative implementation, network performance estimates are directly supported by the network routing service and can be used for replica routing. For example, certain proposed network routing procedures such as IDPR support network routing servers that can determine the expected network performance of a route between two specified IP addresses in an internetwork. This network service can be directly used in steps 535 to 570 or 575 to pick the server replica in the replica routing database with the best expected performance from the server replica's IP address to the client's IP address.

In an alternative implementation, every root replica router runs on the same computer as a server replica. In this implementation, when no server replica can be found for a particular client's IP address and network location, the replica router directly returns the requested information from its local server replica instead of redirecting the client.

The flowchart in Fig. 5a-5b shows how replica routing can be accomplished with client applets. In step 605 a user activates a link that describes a client applet that mediates access to the target behind the link. In step 610 the applet performs internetwork performance discovery as described in Fig. 6, and at step 615 the applet constructs and sends a replica routing request 620 to one or more root replica routers. The request constructed at step 615 includes the network performance table computed at step 610. Once a client performs step 610, this step does not need to be repeated for a configuration-set interval.

At step 630 a replica router uses the source IP

address of message 620, and performs steps 535 to 570 or 575 from Fig. 4a to compute a set of candidate target IP addresses. In the place of step 562, the network performance table computed in step 610 and transmitted in message 620 is used to create an ordered list of network numbers reachable from the client in descending order of performance. These network numbers are looked up in the replica summary records in the replica routing database (step 565), and matching replica summary record entries result in an aggregate performance number being computed (step 570) from both the entry and the network performance provided by client message 620. The top N entries are used as candidate target IP addresses. Each entry in the candidate target IP address list includes a descriptor indicating whether it is a replica router or server replica (this information is determined from the entry's replica advertisement). If none of the networks in the network map in message 620 match replica summary records, then a default set of server replicas is used for the candidate target addresses (step 575).

Once the candidate target IP address list is computed at step 630 a reply message 650 is constructed at step 645 that includes the IP address list, performance metric values corresponding to each element of the list, an indication if each element of the list is a replica router or a server replica, a timestamp, and a service-specific digital signature, and at step 655 the digital signature is authenticated. At step 660, if the candidate IP addresses returned contain replica routers, then control returns to step 615 and the list of candidate IP addresses is expanded and sorted by performance metric value. Otherwise, at step 670 service requests corresponding to the original user action at step 605 are sent in message 675 to a fixed number of the server replicas having the best performance metrics. At step

680 a server replica processes the request, and in the page returned to the client inserts encoded relative links that will lead the client back, for subsequent requests, to the replica that provided the returned page. Reply messages 685 are received by the applet at step 690, and the earliest reply received is selected for processing in accordance with the original user action at step 605.

In an alternative embodiment, the client applet stores a list of all of the servers and replica routers offered in message 650, and, at step 670, simply constructs a single service request to the server replica or replica router having the best aggregate network performance metric value on the list. In the event that a server or replica router does not respond, the applet will return to the saved list to pick another server address or replica router address to try.

In yet another alternative embodiment, at step 510 or 615 the client constructs a replica routing request that is addressed to a predefined broadcast or multicast address. In this embodiment, replica routers listen for a broadcast or multicast request on this address at 515 or 620. Because multiple replica routers can respond to a broadcast or multicast request, the client can pick the first response, or the response that offers the server replica with the best estimated performance from the client's internetwork location.

Although a system has been described in which replica routers and server replicas all implement a single service (e.g., a single collection of information), generalizations to allow replica routers to function for multiple services can be made by one skilled in the art by introducing appropriate unique service identifiers in messages and database entries and modifying the logic above to include service identifiers.

Novel and improved apparatus and techniques for replica routing have been described herein. It is evident

that those skilled in the art may now make numerous uses and modifications of and departures from the specific embodiment described herein without departing from the inventive concept. Consequently, the invention is to be construed as embracing each and every novel feature and novel combination of features present in or possessed by the apparatus and technique herein disclosed and limited solely by the spirit and scope of the appended claims.

What is claimed is:

1. An internetwork replica routing system comprising:

a plurality of server replicas, at least one replica router, and at least one client computer interconnected by a communications internetwork;

the client computer being programmed to cause a network request for access to a server replica to be transmitted over the communications internetwork;

at least one replica router being programmed to receive the network request and to calculate a performance metric value for each of at least some of the server replicas that specifies estimated communication performance between the client computer and the server replica, based upon the client computer's location in the internetwork, and being programmed to direct the client computer to at least one server replica that is estimated to provide good performance based upon the client computer's location in the internetwork, the replica router selecting the server replica to which it directs the client computer based on the performance metric values of the server replicas as calculated by the replica router;

the server replica to which the client computer is directed by the replica router being programmed to respond to the network request from the client computer.

2. The system of claim 1 wherein:

the server replicas are programmed to cause server replica advertisements to be sent to the replica router, the advertisements containing information from which the replica router can calculate the performance metric value; and

the replica router is programmed to maintain a database of the server replica advertisements.

3. The system of claim 2 wherein the replica router is programmed to match the replica advertisements to their

actual source IP address to determine whether any of the server replicas are located behind firewalls.

4. The system of claim 1 wherein:

the client computer is programmed to send a description of its network environment to the replica router; and

the replica router is programmed to calculate the performance metric value for a server replica based upon the description of the client computer's network environment.

5. The system of claim 1 wherein the replica router is programmed to calculate the performance metric value of a server replica based upon the performance metric value of at least one network router located in a path from the client computer to the replica router.

6. The system of claim 1, wherein the client computer is programmed to cause the network request for access to the server replica to be sent to the replica router by multicasting or broadcasting the replica routing request over the communications internetwork.

7. The system of claim 1, wherein there are a plurality of replica routers arranged in a hierarchy, at least one of the replica routers being programmed to direct the client computer to a server replica that is estimated to provide good performance based upon the client computer's location in the internetwork by directing the client computer to a replica router lower in the hierarchy.

8. The system of claim 7 wherein:

at least one the replica routers is programmed to

cause a replica router advertisement to be sent to a replica router higher in the hierarchy, the replica router advertisement containing information from which the replica router higher in the hierarchy can calculate the performance metric value; and

the replica router higher in the hierarchy is programmed to store the replica router advertisement in the database of advertisements.

9. The system of claim 8 wherein the replica router higher in the hierarchy is programmed to match the replica router advertisement to its actual source IP address to determine whether the replica router that caused the replica router advertisement to be sent is located behind a firewall.

10. A method of replica routing in a communications internetwork comprising a plurality of server replicas, at least one replica router, and at least one client computer, comprising the steps of:

causing a network request for access to a server replica to be transmitted from the client computer over the communications internetwork;

receiving the network request at at least one replica router;

calculating, at the replica router, a performance metric value for each of at least some of the server replicas that specifies estimated communication performance between the client computer and the server replica, based upon the client computer's location in the internetwork;

directing the client computer to at least one server replica that is estimated to provide good performance based upon the client computer's location in the internetwork, the server replica to which the client computer is directed being

selected based on the performance metric values of the server replicas as calculated by the replica router;

responding, at the server replica to which the client computer is directed, to the network request from the client computer.

11. The method of claim 10 further comprising the steps of:

causing server replica advertisements to be sent from the server replicas to the replica router, the advertisements containing information from which the replica router can calculate the performance metric values; and

the replica router maintaining a database of the server replica advertisements.

12. The method of claim 11 further comprising the step of matching, at the replica router, the replica advertisements to their actual source IP address to determine whether any of the server replicas or replica routers are located behind firewalls.

13. The method of claim 10 further comprising the steps of:

sending, from the client computer, a description of its network environment to the replica router; and

calculating, at the replica router, the performance metric value for a server replica based upon the description of the client computer's network environment.

14. The method of claim 10 further comprising the step of calculating, at the replica router, the performance metric value of a server replica based upon the performance metric value of at least one network router located in a path

from the client computer to the replica router.

15. The method of claim 10, further comprising the step of causing the network request for access to the server replica to be sent from the client computer to the replica router by multicasting or broadcasting the replica routing request over the communications internetwork.

16. The method of claim 10, wherein there are a plurality of replica routers arranged in a hierarchy, and the method further comprises the step of at least one of the replica routers directing the client computer to a server replica that is estimated to provide good performance based upon the client computer's location in the internetwork by directing the client computer to a replica router lower in the hierarchy.

17. The method of claim 16 further comprising the steps of:

causing a replica router advertisement to be sent from one of the replica routers to a replica router higher in the hierarchy, the replica router advertisement containing information from which the replica router higher in the hierarchy can calculate the performance metric value; and

the replica router higher in the hierarchy storing the replica router advertisement in the database of advertisements.

18. The method of claim 17 further comprising the step of matching, at the replica router higher in the hierarchy, the replica router advertisement to its actual source IP address to determine whether the replica router that caused the replica router advertisement to be sent is located

behind a firewall.

19. An internetwork replica routing system comprising:

a plurality of server replicas, at least one replica router, and at least one client computer interconnected by a communications internetwork;

the server replicas being programmed to cause server replica advertisements to be sent to the replica router, each of the advertisements containing at least one identifier of a network in the communications internetwork to be serviced by the server replica;

the client computer being programmed to cause a network request for access to a server replica to be transmitted over the communications internetwork;

at least one replica router being programmed to maintain a database of the server replica advertisements, to receive the network request from the client computer, and to direct the client computer to at least one of the server replicas based upon the relationship between the networks identified in the advertisements in the database and a network in which the client computer is located;

the server replica to which the client computer is directed by the replica router being programmed to respond to the network request from the client computer.

20. A method of replica routing in a communications internetwork comprising a plurality of server replicas, at least one replica router, and at least one client computer, comprising the steps of:

causing server replica advertisements from the server replicas to be sent to the replica router, each of the advertisements containing at least one identifier of a network

in the communications internetwork to be serviced by the server replica;

causing a network request for access to a server replica to be transmitted from the client computer over the communications internetwork;

the replica router maintaining a database of the server replica advertisements;

receiving the network request from the client computer at the replica router;

directing the client computer to at least one of the server replicas based upon the relationship between the networks identified in the advertisements in the database and a network in which the client computer is located; and

responding, at the server replica to which the client computer is directed by the replica router, to the network request from the client computer.

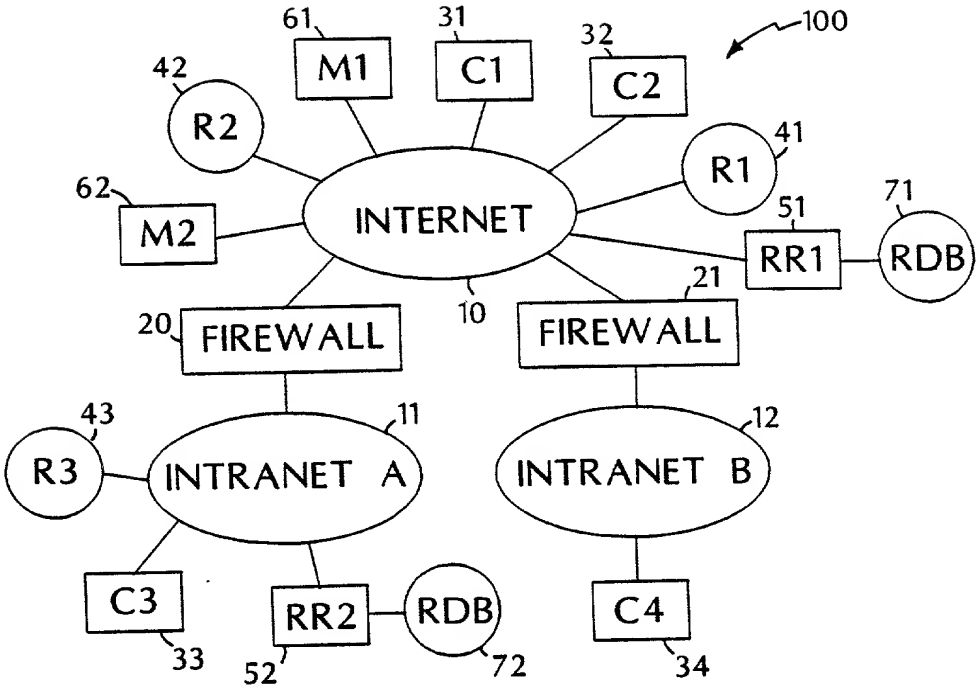
REPLICA ROUTING

Abstract of the Disclosure

The present invention is a new method called replica routing that automatically directs client computers that request a service to a server replica for that service. The replica chosen by replica routing is the replica that is expected to provide the best performance to the client based upon the client's location in the internetwork topology and the estimated performance of the internetwork. In addition, the system and method is designed to permit new replicas to be flexibly added without undue administrative overhead.

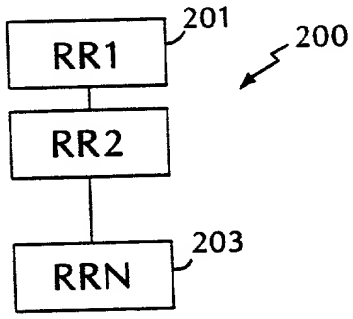
220822.B11

DO NOT ENTER



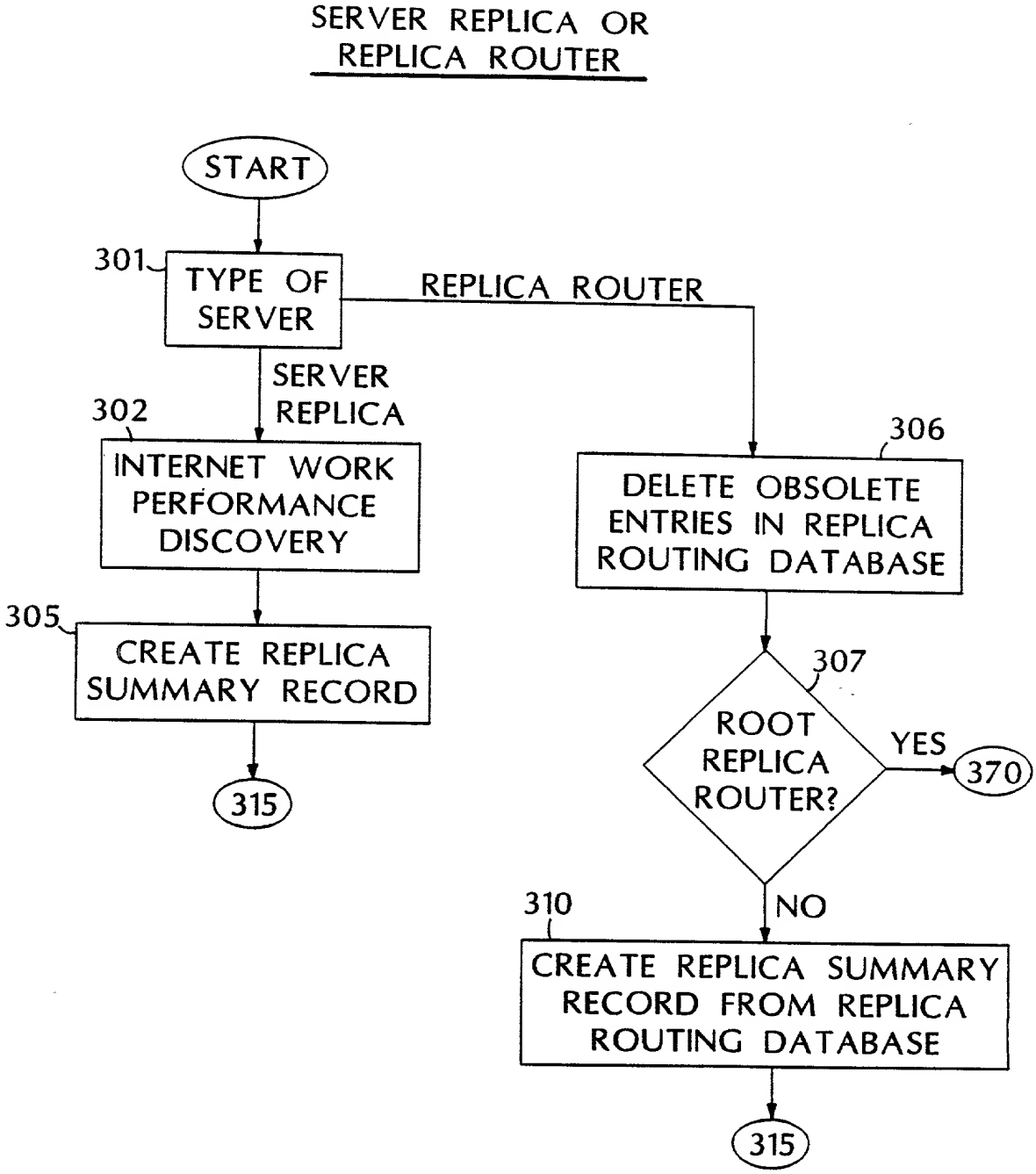
REPLICA ROUTING SYSTEM

FIG. 1



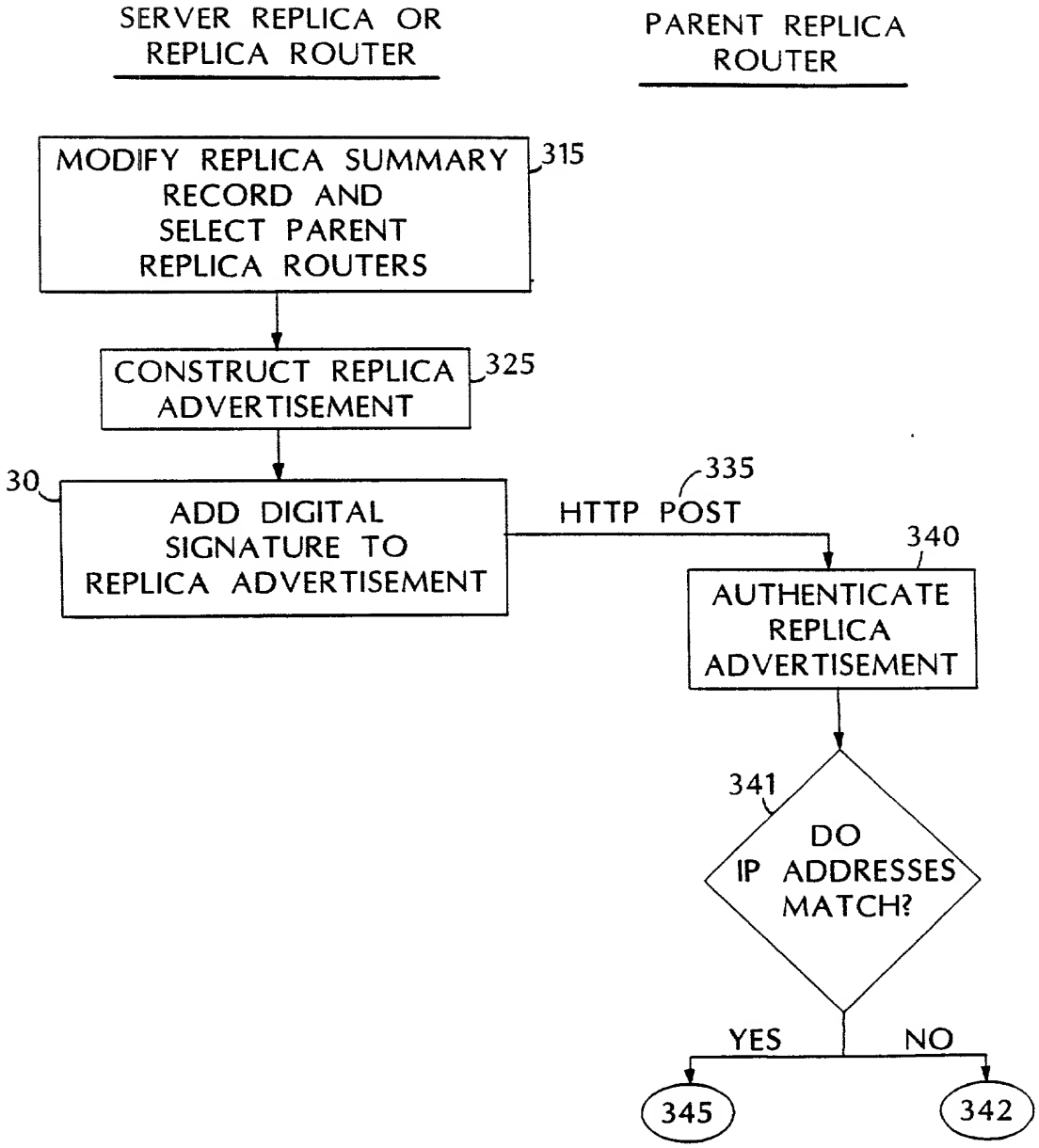
REPLICA ROUTING HIEARCHY

FIG. 2



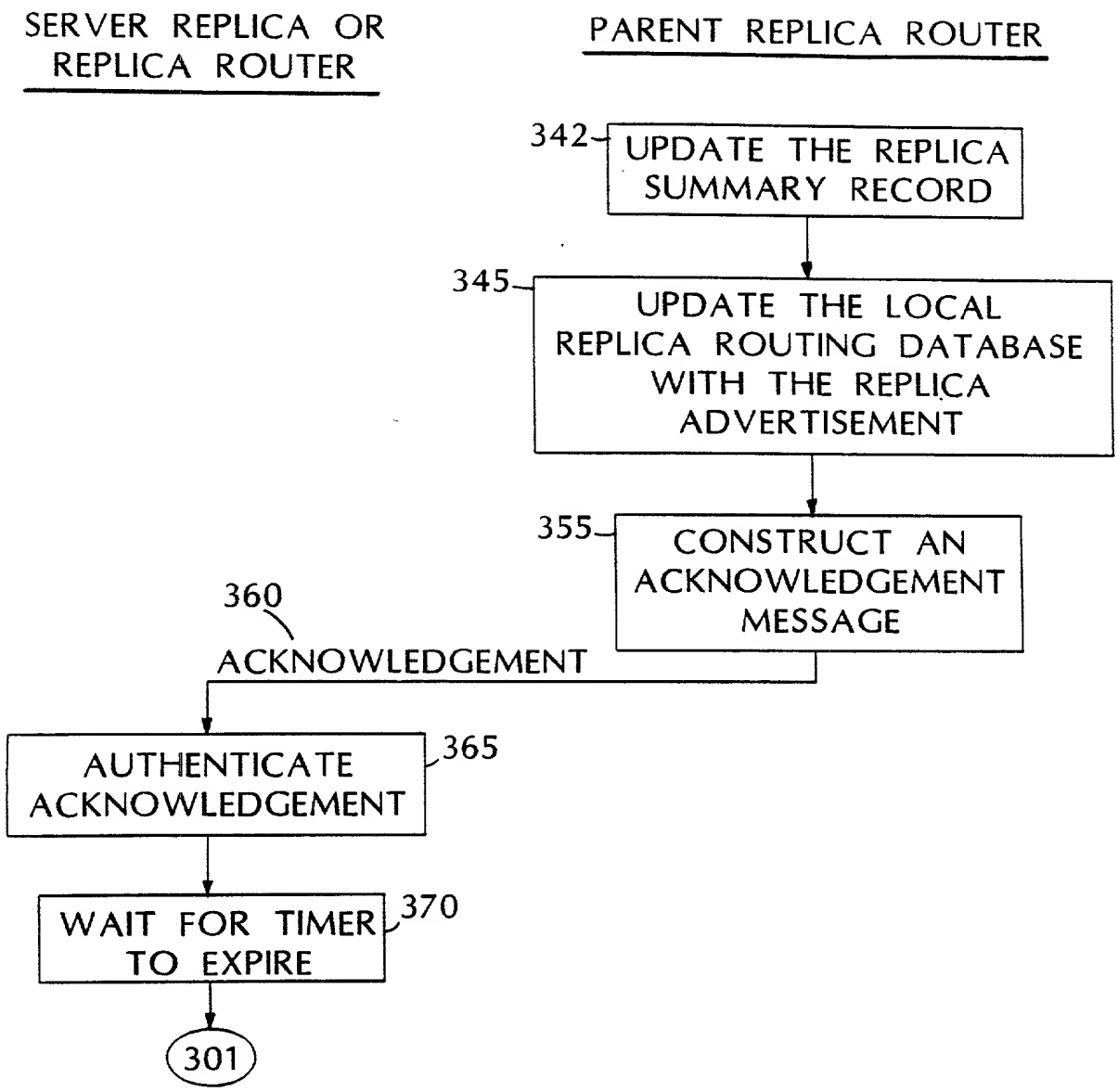
REPLICA ADVERTISEMENT REGISTRATION

FIG. 3



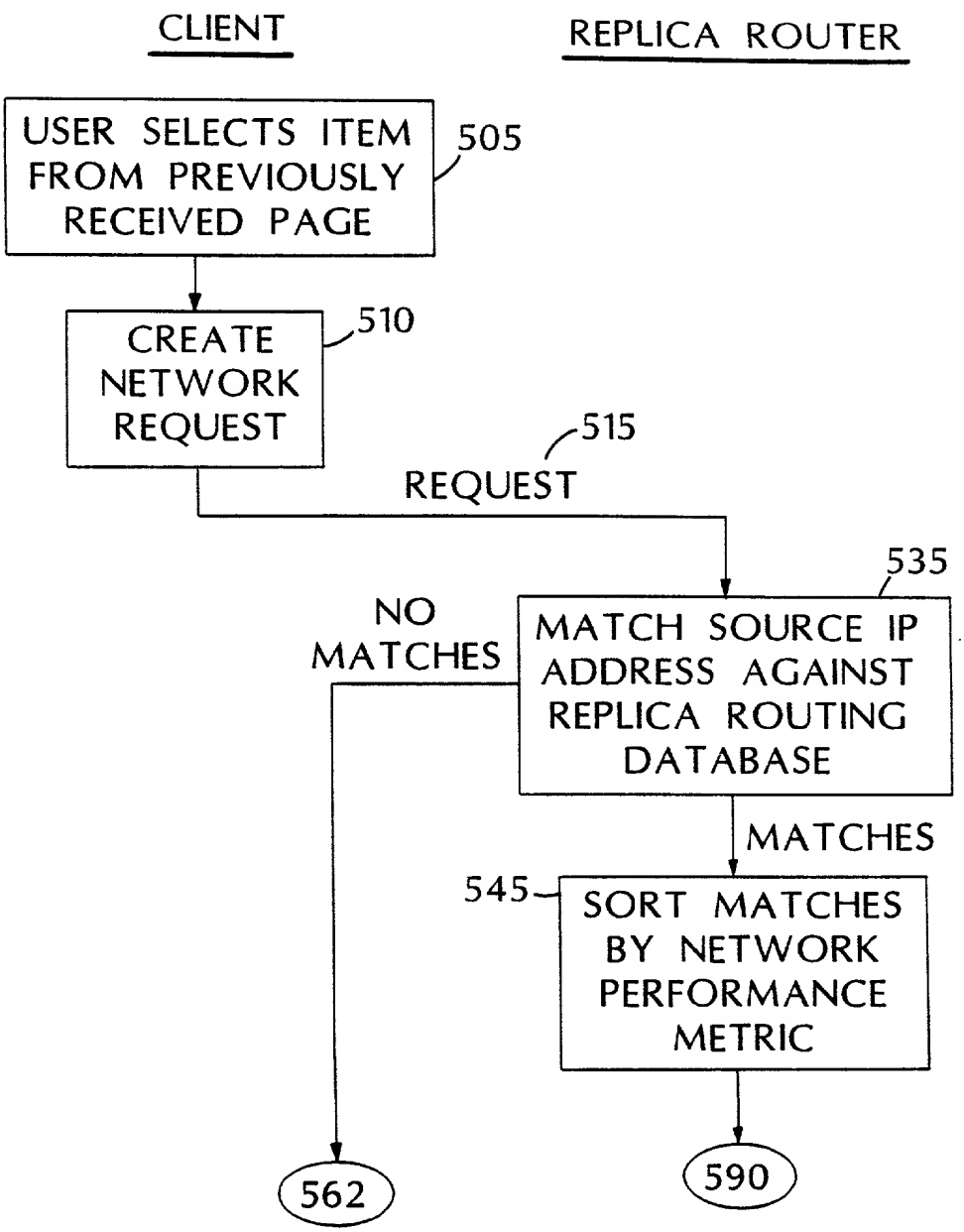
REPLICA ADVERTISEMENT REGISTRATION

FIG. 3B



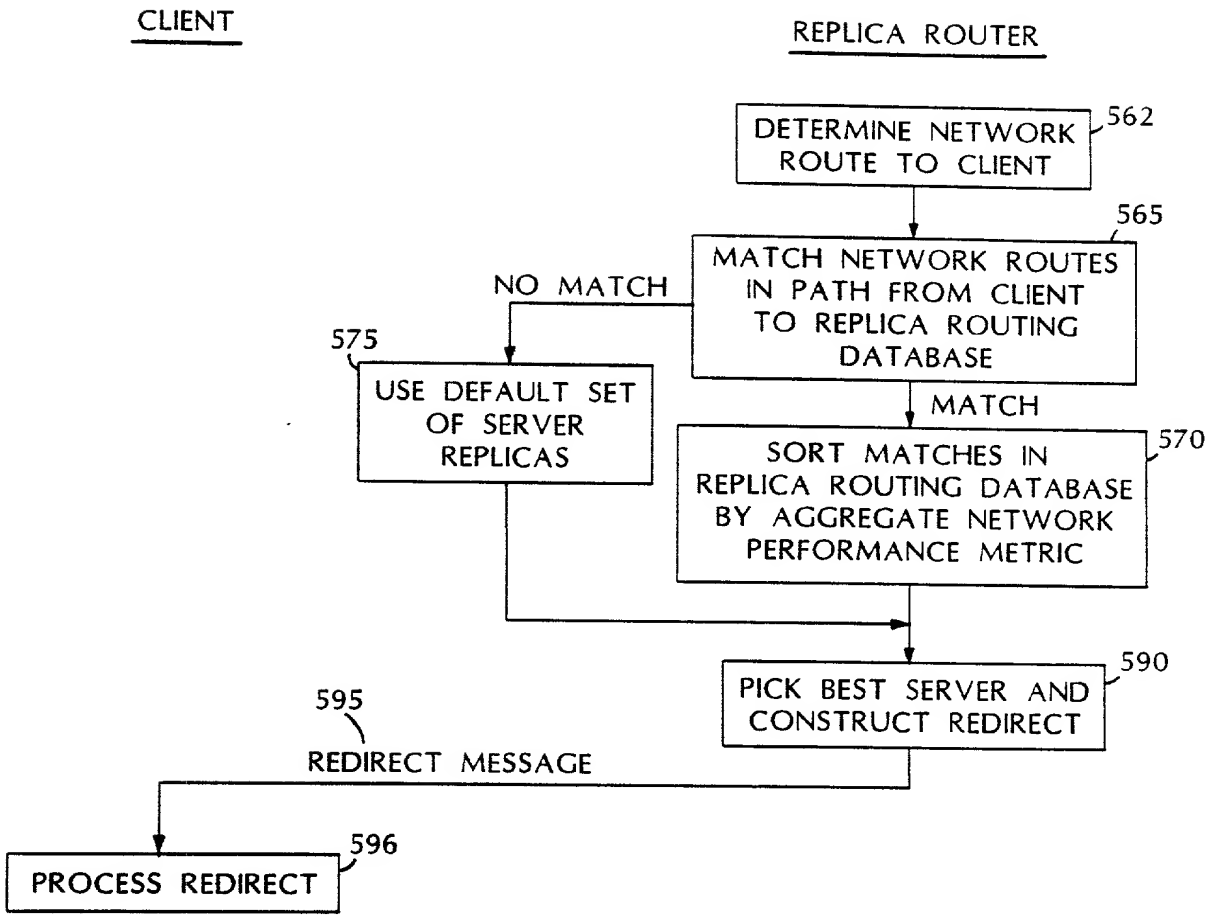
REPLICA ADVERTISEMENT REGISTRATION

FIG. 3C



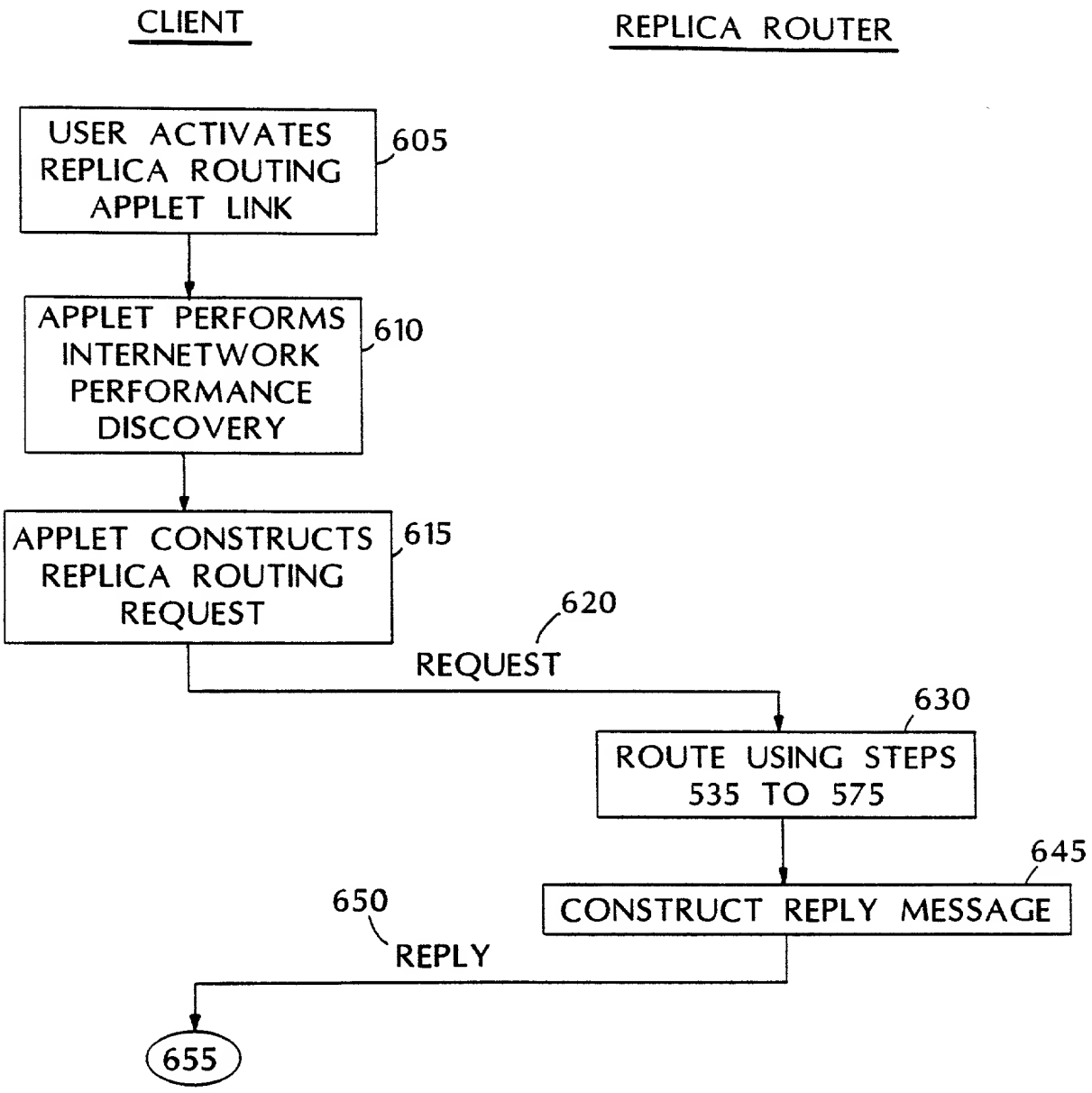
REPLICA ROUTING WITH REDIRECTS

FIG. 4A



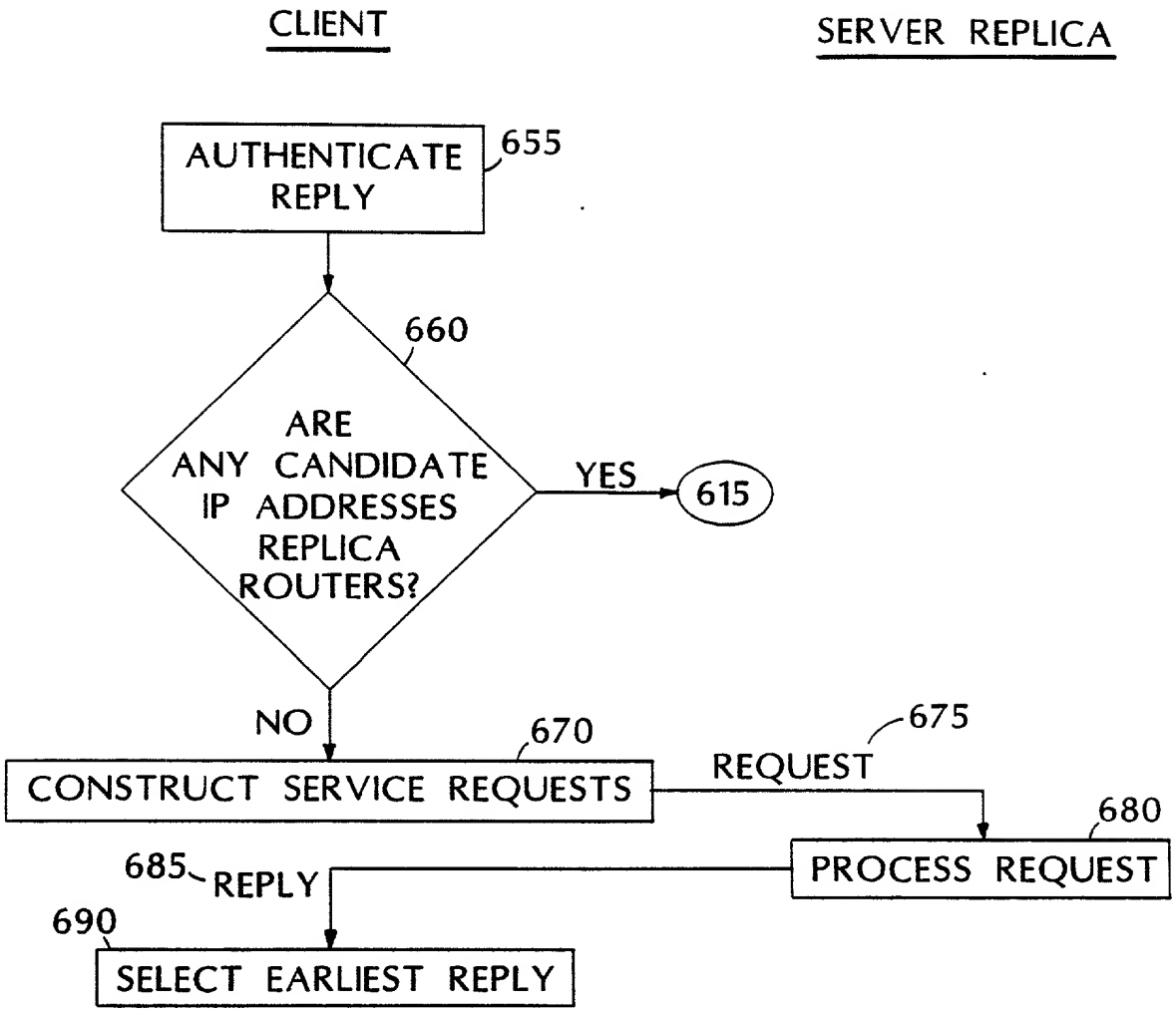
REPLICA ROUTING WITH REDIRECTS

FIG. 4B



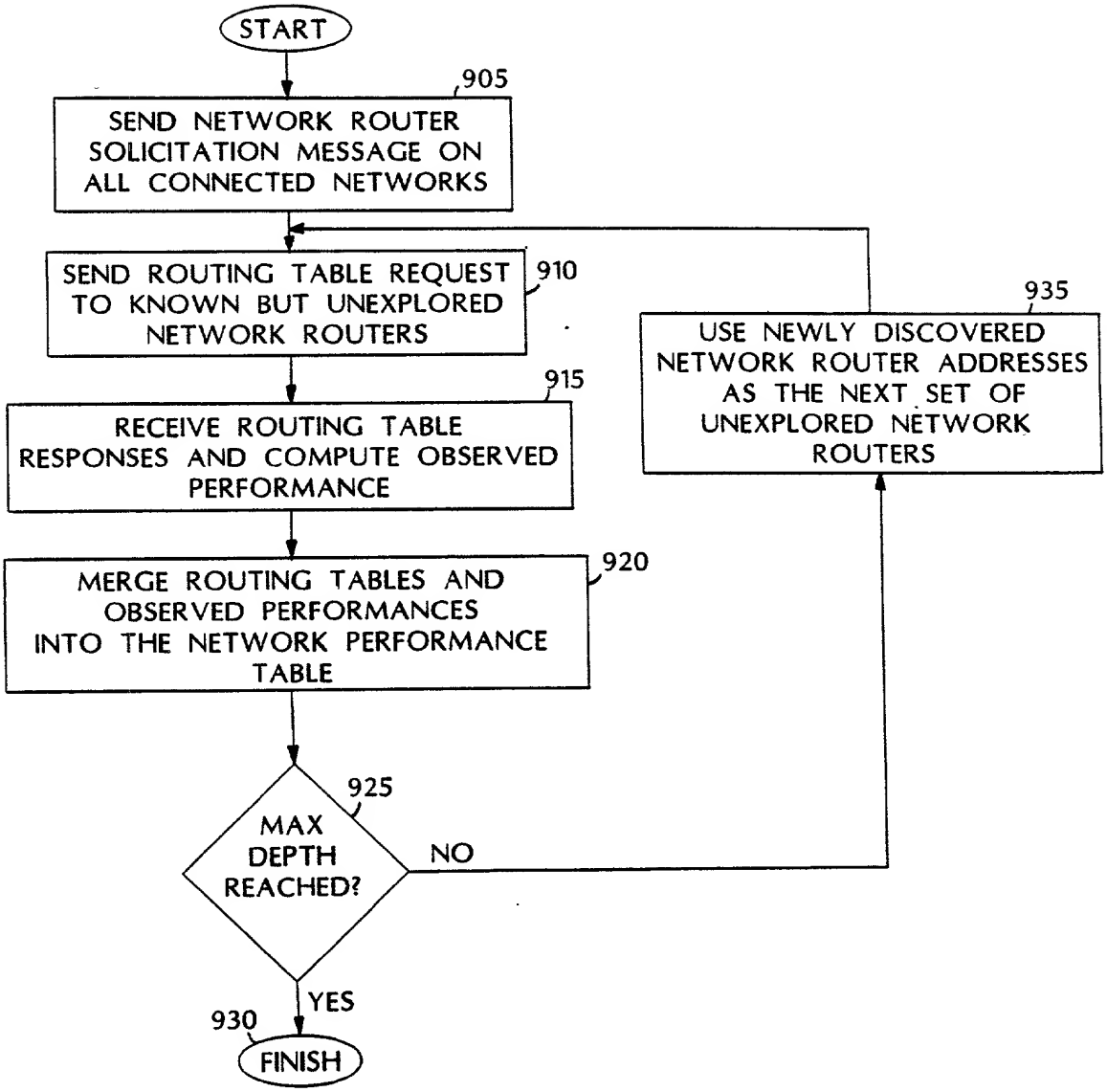
REPLICA ROUTING WITH CLIENT APPLETS

FIG. 5A



REPLICA ROUTING WITH CLIENT APPLETS

FIG. 5B



INTERNETWORK PERFORMANCE DISCOVERY

FIG. 6

PATENT
ATTORNEY DOCKET NO: 06105/003001

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled REPLICA ROUTING, the specification of which

- ■ is attached hereto.
- ☐ was filed on _____ as Application Serial No. _____ and was amended on _____.
- ☐ was described and claimed in PCT International Application No. _____ filed on _____ and as amended under PCT Article 19 on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Gary A. Walpert, Reg. No. 26,098, James E. Mrose, Reg. No. 33,264, John N. Williams, Reg. No. 18,948, Charles C. Winchester, Reg. No. 21,040, Timothy A. French, Reg. No. 30,175.

Address all telephone calls to James E. Mrose at telephone number 617/542-5070.

Address all correspondence to Gary A. Walpert, Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Full Name of Inventor: David K. Gifford

Inventor's Signature: _____ Date: _____

Residence Address: 26 Pigeon Hill Road, Weston, Massachusetts 02193

Citizen of: U.S.A.

Post Office Address: same

PATENT
ATTORNEY DOCKET NO: 06105/003001

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled REPLICA ROUTING, the specification of which

- ☐ is attached hereto.
- ■ was filed on January 7, 1997 as Application Serial No. 08/779,770 and was amended on _____.
- ☐ was described and claimed in PCT International Application No. _____ filed on _____ and as amended under PCT Article 19 on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Gary A. Walpert, Reg. No. 26,098, James E. Mrose, Reg. No. 33,264, John N. Williams, Reg. No. 18,948, Charles C. Winchester, Reg. No. 21,040, Timothy A. French, Reg. No. 30,175.

Address all telephone calls to James E. Mrose at telephone number 617/542-5070.

Address all correspondence to Gary A. Walpert, Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Full Name of Inventor: David K. Gifford

Inventor's Signature: _____ Date: 5/12/97

Residence Address: 26 Pigeon Hill Road, Weston, Massachusetts 02193

Citizen of: U.S.A.

Post Office Address: same

221254 B11